

Mestrado em Estatística e Gestão de Informação
Master Program in Statistics and Information Management

PROVISÃO DE SINISTROS

Estudo de uma Companhia de Seguros e o Impacto
do Fator Cauda

Beatriz Sáinz Dinis Raposo Martinho

Dissertação apresentada como requisito parcial para
obtenção do grau de Mestre em Estatística e Gestão de
Informação

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

PROVISÃO DE SINISTROS: ESTUDO DE UMA COMPANHIA DE SEGUROS E O IMPACTO DO FATOR CAUDA

por

Beatriz Sáinz Dinis Raposo Martinho

Dissertação apresentada como requisito parcial para a obtenção do grau de Mestre em Estatística e Gestão de Informação, Especialização em Análise e Gestão de Risco

Orientadora/Coorientadora: *Professora Doutora Maria de Lourdes Belchior Afonso*

Janeiro 2021

AGRADECIMENTOS

Em primeiro lugar quero agradecer à Professora Doutora Maria de Lourdes Belchior Afonso pela sua valiosa orientação, sugestões, conselhos e pela sua sempre pronta disponibilidade.

Gostaria de agradecer também ao Diretor do departamento para o qual trabalho, Edgar Lopes, por desde o início me ter dado todo o apoio imprescindível para a realização da minha Tese de Mestrado. Agradeço igualmente ao meu colega Paulo Pinheiro pela disponibilização dos dados solicitados, sendo esta a parte fundamental para o sucesso do meu trabalho.

À minha família, em especial aos meus pais e à minha irmã, por todo apoio, paciência e por estarem sempre presentes em todas as etapas da minha vida.

Aos meus amigos mais antigos, a Mariana, a Rita, a Márcia e o Zé, que me acompanham há mais de 15 anos, um grupo para todos os momentos independentemente do percurso que cada um seguiu.

Aos amigos que fui fazendo durante o meu percurso, em especial à Raquel e à Patrícia, por me terem apoiado, motivado e estarem sempre do meu lado.

Por fim, agradecer ao meu grupo de amigos da faculdade, por terem feito com que a minha licenciatura fosse ainda melhor, levando-os para a vida seja no âmbito pessoal como profissional.

RESUMO

O ajuste das provisões técnicas é um tema que tem sido sempre objeto de estudo por parte das Seguradoras. No entanto, existe uma constante preocupação em evitar a ruína, o que leva as Companhias de Seguros a garantir uma gestão de provisões técnicas apropriada e adequada para assegurar as responsabilidades com os seus clientes.

Dos vários tipos de provisões técnicas, as provisões para sinistros são as que têm um impacto mais significativo. Para a sua determinação, existem vários métodos, como os Determinísticos e os Estocásticos. O método Determinístico de *Chain Ladder* é usado para estimar as provisões para sinistros. Para evitar as limitações que essa metodologia apresenta, são utilizados modelos Estocásticos, como o proposto por *Thomas Mack*.

A presente Dissertação pretende avaliar diferentes metodologias de extrapolar o fator cauda no cálculo de provisões para sinistros em ramos com desenvolvimento longo. Aplica-se o modelo *Thomas Mack* (baseado em *Chain Ladder*) para o cálculo das reservas através da informação providenciada por uma Companhia de Seguros na linha de negócio automóvel. Para extrapolar o fator cauda utilizou-se o valor provisionado caso a caso pela Companhia de Seguros, um modelo de ajuste Log-Linear e o método apresentado por *Walter Neuhaus*.

PALAVRAS-CHAVE

Provisão para Sinistros; *Thomas Mack*; Solvência II; *Ultimate*; Fator Cauda.

ABSTRACT

The adequacy of technical provisions is a subject that is studied by the Insurers themselves. There is a fear of failure that incentivizes Insurance Companies to guarantee a proper and skilled management of adequate technical provisions, to ensure the responsibilities they have towards their customers.

Among the various types of provisions, the provisions for claims are the ones that have the most significant impact on the technical provisions. There are two methods, deterministic and stochastic, for its determination. For this purpose, the deterministic method of *Chain Ladder* is usually used to estimate the provisions for claims. However, to go around the limitations of this methodology, stochastic models take place, such as the model proposed by Thomas Mack.

The present work intends to evaluate different methodologies in the extrapolation of the tail factor in the calculation of provisions for claims in the long development branches. The *Thomas Mack* model is applied (based on *Chain Ladder*) for the calculation of provisions for claims on the automobile line of business with the information given by an Insurance Company. To extrapolate the tail factor, a case by case amount provided by the Insurance Company was used, as well as a Log-Linear adjustment model and a Method presented by *Walther Neuhaus*.

KEYWORDS

Provision for Claims; *Thomas Mack*; Solvency II; Ultimate; Tail Factor.

ÍNDICE

Índice de Figuras.....	V
Índice de Tabelas.....	VI
Lista de Siglas e Abreviaturas	VII
1. Introdução.....	1
2. Revisão da Literatura	4
3. Modelo Chain-Ladder	6
4. Modelo Thomas Mack	9
5. Cálculo dos Fluxos de Caixa e o Desdobramento do Fator Cauda	14
6. Aplicação do Modelo Thomas Mack.....	16
7. Resultados e Discussão	22
7.1. Caso 1.....	22
7.2. Caso 2.....	23
7.3. Caso 3.....	25
7.4. Caso 4.....	26
8. Conclusões	31
9. Bibliografia	32

ÍNDICE DE FIGURAS

Figura 3.1 - Formatos de Matrizes de Dados (A), (B) e (C)	8
Figura 6.1 - Montantes Pagos Acumulados (<i>R Cran</i>)	17
Figura 6.2 - Ajustamento dos dados aos fatores de desenvolvimento	17
Figura 6.3 - Resíduos Ponderados	20
Figura 7.1- Extrapolação Log-Linear dos fatores de idade para idade (C1)	25
Figura 7.2 - Padrão do Desenvolvimento Esperado dos Sinistros (C3)	25
Figura 7.3 - Novos Fatores de Desenvolvimento (C4)	26

ÍNDICE DE TABELAS

Tabela 3.1 - Matriz de Dados (Triângulo de Sinistros)	6
Tabela 6.1 - Montantes Incrementais Pagos (em milhares de euros)	16
Tabela 6.2 - Montantes Pagos Acumulados (em milhares de euros).....	16
Tabela 6.3 - Fatores de Desenvolvimento	17
Tabela 6.4 - Coeficientes dos fatores de desenvolvimento Individuais	18
Tabela 6.5 - (A): Matriz ri, k e (B): Matriz si, k	18
Tabela 6.6 - Coeficientes de Correlação de Spearman.....	19
Tabela 6.7 - Conjuntos P e G.....	19
Tabela 6.8 - Momentos das Z_k e Z	20
Tabela 7.1 - <i>Lag Factor</i> (%)	22
Tabela 7.2 - Matriz Incremental preenchida (C1).....	22
Tabela 7.3 - Fluxos de Caixa (C1)	23
Tabela 7.4 - Aplicação do método de Chain Ladder (C1)	23
Tabela 7.5 - Novos Fatores de Desenvolvimento (C2)	23
Tabela 7.6 - Cálculo de Provisão Futura (C2).....	24
Tabela 7.7 - Matriz Incremental Preenchida (C2).....	24
Tabela 7.8 - Fluxo de Caixa (C2).....	25
Tabela 7.9 - Novos Fatores de Desenvolvimento (C3)	25
Tabela 7.10 - Matriz Incremental Preenchida (C3).....	26
Tabela 7.11 - Fluxos de Caixa (C3)	26
Tabela 7.12 - Novos Fatores de Desenvolvimento (C4)	27
Tabela 7.13 - Matriz Incremental Preenchida (C4).....	27
Tabela 7.14 - Fluxos de Caixa (C4)	27
Tabela 7.15 - Resultados das Reservas Obtidas	28

LISTA DE SIGLAS E ABREVIATURAS

ASF	Autoridade de Supervisão de Seguro e Fundos de Pensões
EQM	Erro Quadrático Médio
IBNER	<i>Incurred But Not Enough Reported</i>
IBNR	<i>Incurred But Not Reported</i>
IC	Intervalo de Confiança

1. INTRODUÇÃO

Um contrato de Seguro é um acordo entre duas partes: a seguradora e o segurado. O contrato consiste no fornecimento de um serviço por parte da Seguradora que visa uma troca, por uma pessoa singular ou coletiva, de um risco que pode ser seu ou de terceiros. Assim sendo, o segurado paga à seguradora uma certa taxa (prémio) e, em retorno, é paga uma compensação apropriada pela seguradora no caso da ocorrência de um sinistro.

Pode afirmar-se que as seguradoras têm um processo de negócio invertido daquele a que estamos habituados. O cliente paga o prémio de seguro cobrado pela seguradora e esse prémio é liquidado com base na expectativa de ocorrência de sinistros futuros (valor pago para compensar o cliente, no caso de ocorrência do evento ao segurado).

É neste contexto que aparece a problemática em estudo.

Os custos decorrentes de sinistros podem levar anos a serem reportados pelos segurados às seguradoras e, em alguns casos, a serem totalmente pagos, revelando-se assim mais tarde com valores desconhecidos. Isto deve-se a várias razões: o atraso no reporte da ocorrência do sinistro da parte do segurado à seguradora; o tempo gasto na liquidação de diversos pagamentos, como por exemplo, ao considerar-se as despesas médicas ou os pagamentos dos sinistros derivados a danos pessoais ou pensões, ou quando um caso é levado a tribunal.

Atualmente, as seguradoras devem garantir uma gestão própria e qualificada de provisões técnicas, de forma a assegurar as responsabilidades que têm para com os seus clientes. Assim, as Seguradoras precisam de antecipar o valor total dos sinistros a pagar, dada a sua importância de forma a garantir o pagamento dos seus encargos passivos, da qual as provisões técnicas fazem parte, de acordo com os processos pendentes de sinistros.

As seguradoras têm duas preocupações principais com as reservas de sinistros: uma relacionada com as provisões para sinistros em que os custos estão associados a sinistros que ocorreram, mas não foram reportados (IBNR) e às provisões para sinistros em que os processos não foram ainda resolvidos, sendo que os processos permanecem ainda em aberto (IBNER) (Straubs E., 1998).

De acordo com a Circular nº28/2004 (Autoridade de Supervisão de Seguros e Fundos de Pensões (ASF), 2004): *“A provisão para sinistros corresponde ao custo total estimado que a empresa de seguros suportará para regularizar todos os sinistros que tenham ocorrido até ao final do exercício, quer tenham sido comunicados ou não, após dedução dos montantes já pagos respeitantes a esses sinistros.”*

A estrutura do balanço destaca a importância da seguradora na realização de provisões técnicas adequadas, sendo que o Passivo inclui as responsabilidades através das provisões (Borginho, 2003). As provisões para sinistros destacam-se por causa do seu peso em comparação com os outros tipos de provisão. Assim, a companhia de Seguros deve ter a capacidade para considerar todos os seus futuros compromissos financeiros. No setor de Seguros, existem fatores como a flutuação de sinistros, reservas mal calculadas e gestão ineficiente, entre outros, que influenciam a solvência.

Através do relatório de Dados Estatísticos de Seguros de 2018 (Autoridade de Supervisão de Seguros e Fundos de Pensões (ASF), 2018), pode constatar-se que, no último ano, o montante de provisões para sinistros nos seguros de Não Vida era de 4 536 962 mil euros, representando 79% das provisões técnicas. Desta forma, pode verificar-se que as provisões para sinistros são o tipo de provisões que têm mais impacto nas provisões técnicas. Portanto, há um reforço na ideia da importância do cálculo das provisões para sinistros.

A Solvência II é uma ferramenta importante para os reguladores conseguirem avaliar em contexto externo a Solvência das companhias de Seguros, tendo em conta três pilares estratégicos: os Requisitos Quantitativos, os Requisitos Qualitativos e a Disciplina de Mercado, que promovem a difusão de informações e a transparência da Companhia de Seguros. Consequentemente, o tratamento das provisões técnicas passa pela discussão de vários métodos que podem ser usados para determinar as melhores estimativas de provisões em Seguros Não-Vida no primeiro pilar.

Assim, um dos objetivos desta é permitir à autoridade supervisora os meios adequados para identificar situações adversas que possam existir nas seguradoras sob a sua responsabilidade e garantir uma melhor gestão de risco, com margens de solvência adequadas para permitir a eliminação de qualquer risco.

Para o cálculo das provisões para sinistros (reservas), existem vários métodos estatísticos que podem ser utilizados. Existem dois grandes grupos de métodos estatísticos: os métodos Determinísticos e os métodos Estocásticos.

Os métodos Determinísticos são os mais antigos, que geralmente consideram os fatores de desenvolvimento que refletem a proporcionalidade entre os últimos anos de informação, sendo utilizados para projetar as informações futuras. Desta forma, estes podem estimar o valor esperado dos custos decorrentes de sinistros pendentes da sua data de avaliação.

O problema é que este método apenas dá uma estimativa pontual. Com a introdução da Solvência II e a necessidade de implementação de medidas de variabilidade, os métodos Estocásticos viram na sua importância acrescida nos últimos anos uma base estatística mais rigorosa, permitindo não só a medição de estimativas de provisão, mas também medidas de erro associadas a essas estimativas, algo que os modelos determinísticos não consideram.

A presente Dissertação tem como objetivo conhecer alguns aspetos do negócio de seguros relacionados com a reserva de sinistros, a fim de permitir que as seguradoras cumpram as responsabilidades assumidas quando o cliente paga o prémio.

Este estudo pretende calcular as provisões para os sinistros de danos corporais na linha de negócio Automóvel, utilizando o modelo estocástico obtido por *Thomas Mack* (1993) nas medidas de variabilidade das Estimativas nas Reservas pelo Modelo de *Chain Ladder* para obter os intervalos de confiança para a estimativa dos valores dos sinistros IBNR e IBNER.

Para responder à necessidade de obter os fluxos de Caixa por ano (Solvência II), irá ser adaptado o Modelo de *Thomas Mack* (1993) para considerar os fatores das perdas futuras. Para tal, será utilizado o valor provisionado caso a caso pela Companhia de Seguros, um modelo de ajustamento Log-Linear de extrapolação e um modelo referido em *Walther Neuhaus* (2008).

Com os resultados obtidos, será feita uma análise e uma comparação entre os vários métodos para extrapolar o fator cauda.

2. REVISÃO DA LITERATURA

Para a estimação de provisões para sinistros, uma das metodologias mais utilizadas é a técnica de *Chain Ladder* por ser um dos métodos determinísticos mais práticos e simples, capaz de produzir resultados próximos da realidade. Segundo *Taylor* (2000), esta metodologia foi desenvolvida por *Harneck* (1966) e assume que os fatores de desenvolvimento se mantêm os mesmos (ou semelhantes) no futuro sem assumir qualquer distribuição probabilística.

Thomas Mack em “*Distribution-Free Calculation of the Standard Error of Chain Ladder Reserve Estimates*” (1993) e em “*The Standard Error of Chain Ladder Reserve Estimates: Recursive Calculation and inclusion of Tail Factor*” (1999) desenvolveu o método Estocástico com base no de *Chain Ladder*. A vantagem deste método está em ser possível deduzir as estimativas do Erro Quadrático Médio associado à Variabilidade dos Resultados, podendo para as estimativas pontuais obter-se os intervalos de confiança.

Existem outros métodos estocásticos, nomeadamente a simulação *Bootstrap* e a teoria da Credibilidade Atuarial. Primeiramente, o método *Bootstrap* é uma técnica de simulação, proposta inicialmente por *Efron, B.* (1979) que é baseada na geração de amostras aleatórias com substituição. Por outro lado, a Teoria de Credibilidade, segundo *Norberg, R.* (1986) é um método desenvolvido em meados do século XX, para cobrar os prémios de risco de um determinado portfólio, para o ano $t+1$, com base nas informações disponíveis no ano t e no mercado global, com características específicas do conjunto de riscos que compõem o portfólio.

Apesar de serem matérias de estudo, a presente dissertação irá focar-se apenas no método de *Thomas Mack* para o cálculo das provisões para Sinistros.

Atendendo ao tempo de desenvolvimento dos pagamentos de sinistros, em particular no caso automóvel, é necessário analisar o fator de cauda. Existem vários métodos que ajudam a projetar e a analisar o fator de cauda que pode ser usado. A Casualty Actuarial Society, no seu grupo de trabalho *The Estimation of Loss Development Tail Factors: A Summary Report Cast Tail Factor Working Party* (*Halpert et al.*, 2013) apresenta três tipos de métodos: Os métodos de *Bondy Type*, os métodos Algébricos e os Métodos de Curvas de Ajustamento.

Segundo *Halpert et al.* (2013) os Métodos do Tipo *Bondy* foram desenvolvidos por *Martin Bondy* durante a década de 60 para estimar os anos de desenvolvimentos adicionais ocorridos após o último ano de desenvolvimento que é conhecido no triângulo de Sinistros de determinada matriz. Além disso, é utilizado como fator cauda a última relação de fatores de desenvolvimento observada. Desde então, surgiram várias alterações a este método, baseadas no fator cauda da última relação entre fatores de desenvolvimento disponível.

Halpert et al. (2013) afirma que os Métodos Algébricos se concentram na relação existente entre os triângulos *run-off* dos montantes pagos. Estes são baseados em cálculos simples, não sendo necessária uma fórmula matemática complexa. Além disso, apenas as informações acerca dos montantes pagos disponíveis são necessárias para esses métodos.

Os atuários observaram que podiam também extrapolar o desenvolvimento da cauda através de um ajustamento da curva para os fatores de desenvolvimento anteriores à cauda. Após esse

ajustamento, a curva ajustada é utilizada para extrapolar os fatores de desenvolvimento adicionais à cauda. Através de *Halpert et al.* (2013), os métodos de ajustamento das curvas estudam a relação que existe entre os fatores de desenvolvimento e os anos de desenvolvimento, sendo esta relação uma suposição para ajustar a curva aos fatores de desenvolvimento. Além disso, existem métodos que ajustam a curva aos montantes pagos e outros métodos que se ajustam aos fatores de desenvolvimento. O que é comum entre estes é que ambos começam com a diminuição dos fatores de desenvolvimento ao longo dos anos, dando origem a um conjunto de curvas. De seguida, são selecionados os coeficientes que especificam o membro do conjunto de curvas que se ajusta aos dados.

Um desses casos é o método utilizado em *Neuhaus, W.* (2008), onde um modelo é apresentado para a estimativa de sinistros pendentes que abrange o atraso em duas dimensões: no reporte e na avaliação de um sinistro. Este modelo traz uma distinção entre o custo dos sinistros reportados e o custo dos sinistros não reportados. Desta forma, o método baseia-se nas teorias de ajuste da curva que visam transmitir o raciocínio dos mecanismos envolvidos na resolução do processo de sinistro.

3. MODELO CHAIN-LADDER

Os dados utilizados para estimar a provisão para sinistros organizados e apresentados na forma de uma matriz incompleta conhecida pelo Triângulo de Sinistros ou o Triângulo *run-off* (Tabela 3.1) (Kärikä M., n.d.).

Na matriz, as quantidades $X_{i,j}$ podem representar o valor total ou a média dos valores pagos das compensações derivadas das responsabilidades pelos sinistros, o número de sinistros reportados, o valor do prémio ou o número de apólices numa Companhia de Seguros. Esta matriz deve ser apurada agregando riscos homogêneos. Para tal, agrupam-se os riscos em coberturas dentro da linha de negócio, como seja no caso em estudo, a cobertura de danos próprios da linha de negócio automóvel.

As colunas da matriz são o período de desenvolvimento dos sinistros, ou seja, o número de períodos que ocorrem até os sinistros serem reportados à seguradora ou os períodos em que ocorre pagamento/reembolso de indemnizações. Estes podem estar divididos em meses, trimestres, anos, entre outros. Para o presente estudo, apenas serão considerados os anos. Por outro lado, cada linha representa o período de ocorrência dos sinistros, ou seja, o seu período de reporte. O período $i=0$ é considerado o primeiro período e o período $i=l$ é considerado o último período. A diagonal é o período em que os pagamentos foram feitos. A coluna de ∞ é a coluna final, que inclui a estimativa dos pagamentos de todos sinistros que ocorreram num período i , mas em que a regularização desses sinistros só será realizada após l períodos. Os dados utilizados irão seguir a hipótese do número de anos de ocorrência igual ao número de anos de desenvolvimento ($i=j$).

Tabela 3.1 - Matriz de Dados (Triângulo de Sinistros)

Período de Ocorrência	Período de Desenvolvimento								
	0	1	2	...	j	...	l-1	l	∞
0	$X_{0,0}$	$X_{0,1}$	$X_{0,2}$...	$X_{0,j}$...	$X_{0,l-1}$	$X_{0,l}$	$X_{0,\infty}$
1	$X_{1,0}$	$X_{1,1}$	$X_{1,2}$...	$X_{1,j}$...	$X_{1,l-1}$		
2	$X_{2,0}$	$X_{2,1}$	$X_{2,2}$...	$X_{2,j}$...			
...			
j	$X_{j,0}$	$X_{j,1}$	$X_{j,2}$...	$X_{j,j}$				
...						
l-1	$X_{l-1,0}$	$X_{l-1,1}$							
l	$X_{l,0}$								

Existem várias componentes que podem ser utilizadas dependendo da escolha do que se deseja estimar. Na presente dissertação, a Tabela 3.1 trata a matriz incremental dos valores pagos, sendo que $X_{i,j}$ são os pagamentos no período de desenvolvimento j , derivado dos sinistros ocorridos no período i . Também pode ser calculada a matriz acumulada.

Assim, $A_{i,k}$ é calculado através da soma dos valores pagos incrementais.

$$A_{i,k} = \sum_{j=0}^k X_{i,j} \quad 3.1$$

A matriz está dividida em dois triângulos: o triângulo superior à esquerda que tem toda a informação já conhecida e o triângulo inferior à direita que é a provisão que se quer estimar. Assim, o objetivo principal é estimar o valor a ser pago no futuro. Com a estimativa, será possível fazer uma previsão da quantidade esperada de custos para cada período. Depois disso, será possível obter o valor da provisão a ser calculado para o valor presente, de forma a lidar com os custos existentes.

Para o cálculo das reservas, pelo método de *Chain Ladder* é necessário obter os fatores de desenvolvimento (f_k) da seguinte forma:

$$\hat{f}_k = \frac{\sum_{i=0}^{I-k-1} A_{i,k+1}}{\sum_{i=0}^{I-k-1} A_{i,k}}, k = 0, \dots, I-1 \quad 3.2$$

Depois de calculados os fatores de desenvolvimento, procede-se à projeção do valor dos sinistros a ser pago no futuro (o triângulo inferior da matriz).

Para se preencher o triângulo inferior utiliza-se a forma recorrente,

$$\hat{A}_{i,k} = A_{i,k-1} \times \hat{f}_{k-1}, i = 1, \dots, I \wedge k = I-i+1, \dots, I \quad 3.3$$

Para efetuar o cálculo, os fatores de desenvolvimento obtidos anteriormente são multiplicados pela última diagonal da matriz acumulada e assim sucessivamente, até o triângulo inferior ser concluído.

$$\hat{A}_{i,k} = A_{i,I-k} \times \hat{f}_{I-k} \times \dots \times \hat{f}_{I-1}, i = 1, \dots, I \wedge k = i+1, \dots, I \quad 3.4$$

Um indicador bastante útil que pode ser obtido à conta dos coeficientes de desenvolvimento é o *Lag Factor*, que se pode obter por:

$$\widehat{L}_k = \frac{1}{\widehat{F}_k}, k = 0, \dots, I-1 \quad 3.5$$

Onde,

$$\widehat{F}_k = \prod_{j=k}^{I-1} \hat{f}_j, k = 0, \dots, I-1 \quad 3.6$$

O *Lag Factor* indica-nos a percentagem de pagamentos acumulados efetuada até determinado ano de desenvolvimento.

Para calcular o valor das reservas estimadas para os anos de ocorrência, subtrai-se o valor dos pagamentos estimados (conhecido também por *ultimate*) com os pagamentos já realizados (conhecidos também pela última diagonal conhecida da matriz).

$$\hat{R}_i = \hat{A}_{i,\infty} - A_{i,I-i}, i=1, \dots, I \quad 3.7$$

Também existem outras formas de dispor o formato das matrizes, onde a escolha da matriz depende dos dados válidos e disponíveis (Borginho, 2001).

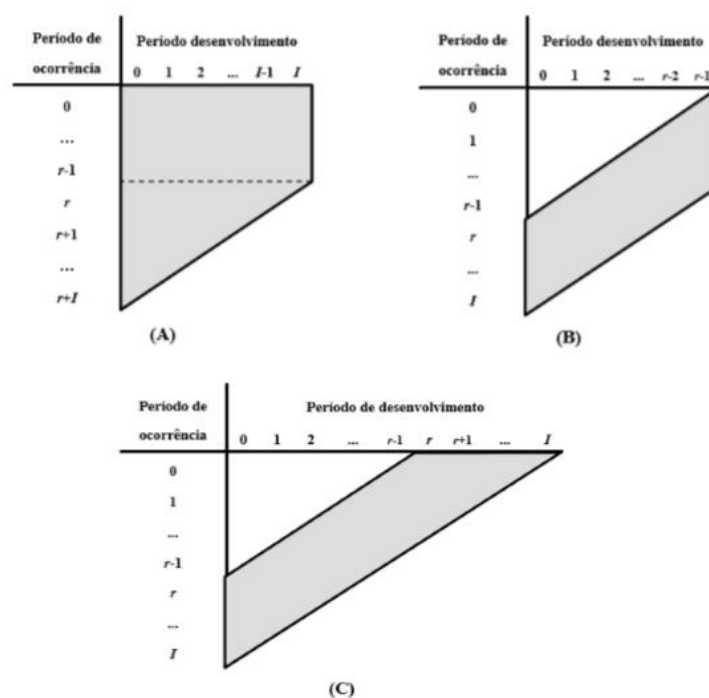


Figura 3.1 - Formatos de Matrizes de Dados (A), (B) e (C)

Fonte: Hugo Borginho (2001)

Um dos formatos possíveis de se obter é o (A), quando a informação é derivada de vários anos de ocorrência. No fim de I anos de Desenvolvimento, a informação pode mostrar-se incompleta, nomeadamente no caso de as observações dos anos de maior desenvolvimento serem consideradas menos fundamentais.

Observando o (B), a hipótese representa o ano de pagamento r , criando um formato diferente dos outros.

Em (C), o formato da matriz apresenta-se semelhante ao de (B), mas com uma extensão dos anos de desenvolvimento onde as observações estão documentadas.

Apesar das alternativas apresentadas nas matrizes, será ainda assim considerado o formato de matriz apresentado na Tabela 3.1, que representa o formato mais comum de ser utilizado neste contexto.

4. MODELO THOMAS MACK

A aplicação do modelo de *Thomas Mack* nos dados de sinistros implica a verificação de algumas hipóteses que precisam de ser testadas, através de algumas técnicas relacionadas com o conjunto de dados em análise. A partir dos resultados desses testes, observam-se as condições para avaliar se o modelo de *Thomas Mack* será apropriado para o conjunto de dados em estudo.

A primeira hipótese constitui na verificação de existência de proporcionalidade entre os anos de desenvolvimento (as colunas da matriz), assumindo que existem fatores de desenvolvimento (Equação 3.2). A Equação 3.4 traduz a relação entre os valores $A_{i,k}$ e o declive dos fatores de desenvolvimento (f_k). Desta forma, ao construir um gráfico com a linha que passa na origem com declive f_k , os pares ordenados $(A_{i,k}, A_{i,k+1})$ deverão estar próximos dessa linha.

Na possibilidade de desvios significativos, devem ser procurados outros métodos que melhor se ajustem aos dados em análise ou, se não for exequível, rejeitar a aplicação do modelo.

É também necessário verificar a hipótese da não correlação entre os fatores de desenvolvimento individuais, uma vez que a hipótese implica que não há uma relação entre os fatores de desenvolvimento individuais. Para testá-la é utilizado o teste de *Spearman*.

O teste de *Spearman* (*Kendall, 1973*) é uma medida não paramétrica utilizada para testar a dependência entre duas variáveis, avaliando através da análise da descrição das variáveis uma função monótona. Este teste tem a principal vantagem de não assumir qualquer distribuição para os dados.

Primeiramente, são calculados os fatores de desenvolvimento individuais com a matriz de valores pagos acumulados.

$$A_{i,k+1}/A_{i,k} \quad 4.1$$

De seguida, são construídas duas novas matrizes, calculadas através dos fatores de desenvolvimento individuais. A primeira matriz são os fatores de desenvolvimento ordenados de ordem crescente ($r_{i,k}$), onde é considerado o ano de fator de desenvolvimento k . A segunda matriz é a de fatores de desenvolvimento individuais ordenados pelos fatores precedentes ($s_{i,k}$), sendo os últimos fatores descartados.

O próximo passo é determinar o coeficiente da correlação do teste de *Spearman* (*Rohatgi, 2015*), que é definido por:

$$T_k = 1 - 6 \times \sum_{i=0}^{I-k-1} \frac{(r_{i,k} - s_{i,k})^2}{[(I-k)^3 - I + k]}, 1 \leq k \leq I-2 \quad 4.2$$

Valores de T_k próximos de 0 indicam que os fatores de desenvolvimento que estão entre os anos $k-1$ e k e os fatores que estão entre k e $k+1$ não são correlacionados. Valores positivos de T_k indicam uma correlação positiva e valores negativos indicam uma correlação negativa.

Se houver ausência de correlação, o valor esperado deve ser $E(T_k) = 0$ e a variância deve ser calculada através de:

$$V(T_k) = \frac{1}{I-k-1} \quad 4.3$$

Com o triângulo de desenvolvimento completo, será ponderada a média dos T_k 's, onde os seus ponderadores serão inversamente proporcionais de $V(T_k)$.

Desta forma, para obter o estimador de variância mínima, calcula-se o T como:

$$T = \frac{\sum_{k=1}^{I-2} ((I-k-1) \times T_k)}{\sum_{k=1}^{I-2} (I-k-1)} \quad 4.4$$

Considerando que a distribuição T_k se aproxima de uma distribuição Normal e tendo em conta que T é obtido através da soma ponderada de vários T_k 's não correlacionados, pode assumir-se que T tem uma distribuição assintótica Normal pela aplicação do Teorema de Limite Central.

Desta forma, se a estimativa de T obtida pelo teste de *Spearman* estiver fora do seguinte intervalo com nível de confiança $100(1 - \alpha)$ por cento, sendo $\alpha \in]0,1[$, rejeita-se a hipótese de ausência de correlação entre os fatores individuais de desenvolvimento.

$$[-\Phi^{-1}(1 - \frac{\alpha}{2})\sqrt{V(t)}; \Phi^{-1}(1 - \frac{\alpha}{2})\sqrt{V(t)}] \quad 4.5$$

, sendo $\Phi^{-1}(1 - \frac{\alpha}{2})$ o quantil da probabilidade de $1 - \frac{\alpha}{2}$ da distribuição Normal standard.

Se T não pertencer ao intervalo de confiança definido acima, as correlações precisam de ser analisadas em maior detalhe e, eventualmente, serem tomadas decisões de não aplicar o modelo presente.

É necessário averiguar também a independência entre os diferentes anos de ocorrência (as linhas da matriz acumulada). A intenção deste teste é verificar se os estimadores dos fatores de desenvolvimento se encontram centrados.

Para realizar este teste é necessário considerar a matriz de fatores de desenvolvimento individuais que são divididos, por ano de desenvolvimento, em dois conjuntos: um com metade dos fatores de desenvolvimento, os fatores grandes (G) e outro com os restantes, os pequenos (P). Deve observar-se se existem diagonais na matriz de dados com a supremacia de elementos de um dos conjuntos.

Para cada diagonal do triângulo (A_k), o número de elementos G e P deve ser contabilizado. Em caso de independência, o número de elementos que pertence ao conjunto G deve ser igual ao número de elementos do conjunto P. Em diagonais que o número total de fatores for ímpar, o valor mediano é retirado. No final, cada fator de desenvolvimento individual pertence a um conjunto ou foi excluído. Assim, há uma probabilidade de 50% de um fator de desenvolvimento individual pertencer ao agrupamento G ou P.

Ao considerar G_k como o número de elementos G que pertencem a A_k e P_k como o número de elementos P que pertencem a A_k , se $Z_k = \min(G_k, P_k)$ for significativamente menor que o valor de $\frac{G_k + P_k}{2}$, haverá uma preponderância de fatores grandes ou pequenos na diagonal k dos fatores de desenvolvimento, onde neste caso a hipótese é rejeitada.

Dito isto, a hipótese não deve ser rejeitada se G_k e P_k tiverem ambos uma Distribuição Binomial com parâmetros $n = G_k + P_k$ e $p = 0,5$. Assumindo uma Distribuição Normal, a hipótese não deve ser rejeitada se $Z = \sum Z_k$, e este estiver dentro do intervalo de confiança calculado pela expressão:

$$[E(Z) - (\Phi^{-1}(1 - \frac{\alpha}{2}) \sqrt{V(Z)}) ; E(Z) + (\Phi^{-1}(1 - \frac{\alpha}{2}) \sqrt{V(Z)})] \quad 4.6$$

O valor esperado e respetiva variância são calculados através das expressões $E(Z) = \sum E(k)$ e $V(Z) = \sum V(k)$, onde cada um dos momentos são calculados por:

$$E(Z_k) = \frac{n}{2} - C_m^{n-1} \cdot \frac{n}{2^n} \quad 4.7$$

$$V(Z_k) = \frac{n(n-1)}{4} - C_m^{n-1} \times \frac{n(n-1)}{2^n} + E(Z_k) - [E(Z_k)]^2 \quad 4.8$$

O m é dado por:

$$m \leq \frac{n(n-1)}{2} \quad 4.9$$

Relativamente à última hipótese, esta consiste em averiguar se os estimadores dos fatores de desenvolvimento obtidos são aqueles que representam menor variância. São construídos gráficos para os ponderadores residuais (desvios entre os dados observados e os dados estimados) deste método e observa-se se o conjunto de pontos obtidos representam algum tipo de tendência. Caso contrário, pode admitir-se que é válido nesse processo.

$$V(A_{i,k-1} | A_{i,0}, \dots, A_{i,k}) = A_{i,k} \cdot \sigma_k^2, \quad 0 \leq i \leq I, \quad 0 \leq k \leq I \quad 4.10$$

Para verificar se é válido, basta observar que a equação acima indica que a variação condicional de $A_{i,k+1}$ é diretamente proporcional com $A_{i,k}$, que possui proporcionalidade constante com σ_k^2 .

Se a hipótese for válida, os valores estimados para a variação condicional serão fechados para $A_{i,k} \cdot \sigma_k^2$. Isto significa que os resíduos obtidos também serão proporcionais ou aproximados a $A_{i,k}$. Desta forma, basta representar os pares ordenados em gráfico.

Se houver alguma tendência no gráfico de resíduos, deve haver uma pesquisa para outros estimadores para \hat{f}_k ou rejeitar a aplicação do presente modelo.

Estimação das provisões e medida de variabilidade

Após a verificação dos três pressupostos, o passo seguinte é estimar as provisões e a sua medida de variabilidade. Para tal, é feita a estimativa do erro padrão (EP) e a construção dos respetivos intervalos de confiança relacionados com os resultados obtidos com o método de *Chain Ladder*.

Para perceber a distância entre a estimativa $\hat{A}_{i,k}$ e o valor que se irá observar $A_{i,k}$, recorre-se ao Erro Quadrático Médio.

$$EQM(\hat{A}_{i,k}) = E[(A_{i,k} - \hat{A}_{i,k})^2 | A] \quad 4.11$$

Onde $A = \{A_{i,k} | i + k \leq I\}$ representa o conjunto de todos os dados até ao momento conhecidos da matriz de triângulo de sinistro.

Para efeitos do cálculo dos EQM, é necessário um estimador centrado, considerando-se:

$$\sigma_k^2 = \frac{1}{I-k-1} \sum_{i=0}^{I-k-1} A_{i,k} \left(\frac{A_{i,k+1}}{A_{i,k}} - \hat{f}_k \right)^2, 0 \leq k \leq I-2 \quad 4.12$$

Dado que a expressão não possibilita o cálculo de um estimador para $\hat{\sigma}_{I-1}^2$, e assumindo que $\hat{\sigma}_{I-3}^2 / \hat{\sigma}_{I-2}^2 = \hat{\sigma}_{I-2}^2 / \hat{\sigma}_{I-1}^2$ se verifica, pelo menos enquanto $\hat{\sigma}_{I-3} > \hat{\sigma}_{I-2}$, utiliza-se a expressão:

$$\hat{\sigma}_{I-1}^2 = \min \left[\frac{\hat{\sigma}_{I-2}^4}{\hat{\sigma}_{I-3}^2}, \min(\hat{\sigma}_{I-3}^2, \hat{\sigma}_{I-2}^2) \right] \quad 4.13$$

Para o cálculo de σ_I^2 utiliza-se uma aproximação para $EQM(\hat{f}_k)$ e $EQM\left(\frac{A_{i,k}}{A_{i,k-1}}\right)$, considerando que se $\hat{f}_{k-1} > \hat{f}_I > \hat{f}_k$, então verifica-se também que:

$$EQM\left(\frac{A_{i,k}}{A_{i,k-1}}\right) > EQM\left(\frac{A_{i,\infty}}{A_{i,I}}\right) > EQM\left(\frac{A_{i,k+1}}{A_{i,k}}\right) \quad 4.14$$

A fórmula usada para esta estimativa será:

$$\hat{\sigma}_N^2 = 1 - \left(\frac{\hat{f}_{k-1} - \hat{f}_I}{\hat{f}_{k-1} - \hat{f}_k} \right) \cdot \hat{\sigma}_{k-1}^2 + \left(\frac{\hat{f}_{k-1} - \hat{f}_I}{\hat{f}_{k-1} - \hat{f}_k} \right) \cdot \hat{\sigma}_k^2 \quad 4.15$$

Se $\hat{f}_{I-1} = 1$ ou $\hat{f}_I = 1$ pode também determinar-se que, $\hat{\sigma}_{I-1}^2 = 0$ e $\hat{\sigma}_I^2 = 0$, respetivamente, sendo que para estes casos espera-se que não haja sinistros a serem pagos com mais de I-1 ou I anos de desenvolvimento.

Para se conhecer a variabilidade dos valores estimados da parte inferior do triângulo, a expressão irá admitir o cálculo do EQM de cada $\hat{A}_{i,k}$ desconhecido e igualmente o $EQM(\hat{A}_{i,\infty})$:

$$EQM(\hat{A}_{i,\infty}) = \hat{A}_{i,k}^2 [EQM\left(\frac{A_{i,k+1}}{A_{i,k}}\right) + EQM(\hat{f}_k)] + EQM(\hat{A}_{i,k}) \cdot \hat{f}_k^2 \quad 4.16$$

Para aplicação desta fórmula, é necessário determinar:

$$EQM\left(\frac{A_{i,k+1}}{A_{i,k}}\right) = \frac{\hat{\sigma}_k^2}{\hat{A}_{i,k}} \quad 4.17$$

Para cada fator de desenvolvimento, a estimativa do Erro Quadrático Médio é dada por:

$$EQM(\hat{f}_k) = \frac{\hat{\sigma}_k^2}{\sum_{i=0}^{I-k-1} A_{i,k}}, 0 \leq i \leq I \text{ e } 0 \leq k \leq I-1 \quad 4.18$$

Partindo do pressuposto que o fator de desenvolvimento \hat{f}_I está entre \hat{f}_{k-1} e \hat{f}_k , o $EQM(\hat{f}_I)$ é estimado através de:

$$EQM(\hat{f}_I) = [EQM(\hat{f}_k) + (EQM(\hat{f}_{k-1}) - EQM(\hat{f}_k)) \cdot \left(\frac{\hat{f}_I - \hat{f}_k}{\hat{f}_{k-1} - \hat{f}_k} \right)] \quad 4.19$$

Apurar o valor de $EQM(\hat{A}_{i,\infty})$ equivale a determinar o valor de $EQM(\hat{R}_i)$:

$$E\hat{Q}M(\hat{R}_i) = E\hat{Q}M(\hat{A}_{i,\infty}) \quad 4.20$$

Por fim, pretende-se calcular a variabilidade do valor da provisão total a constituir.

O estimador pretendido para $E\hat{Q}M(\hat{R})$ será obtido através da seguinte equação:

$$E\hat{Q}M(\hat{R}) = E\hat{Q}M(\sum_{i=0}^I \hat{R}_i) = E\hat{Q}M(\sum_{i=0}^I \hat{A}_{i,\infty}) \quad 4.21$$

Considerando que:

$$\begin{aligned} E\hat{Q}M(\sum_{i=I-k}^I \hat{A}_{i,k+1}) &= \sum_{i=I-k}^I \hat{A}_{i,k+1}^2 \cdot E\hat{Q}M\left(\frac{\hat{A}_{i,k+1}}{\hat{A}_{i,k}}\right) + \\ &(\sum_{i=I-k}^I \hat{A}_{i,k})^2 \cdot E\hat{Q}M(\hat{f}_k) + E\hat{Q}M(\sum_{i=I-k+1}^I \hat{A}_{i,k} \cdot \hat{f}_k^2) \end{aligned} \quad 4.22$$

Neste caso, a estimativa segue uma distribuição assintoticamente normal, obtendo-se os seguintes intervalos de confiança a $100(1 - \alpha)$ por cento para as estimativas das provisões \hat{R}_i e \hat{R} , respetivamente,

$$[\hat{R}_i - \phi^{-1}(1 - \frac{\alpha}{2}) \cdot \sqrt{E\hat{Q}M(\hat{R})}; \hat{R}_i + \phi^{-1}(1 - \frac{\alpha}{2}) \cdot \sqrt{E\hat{Q}M(\hat{R})}] \quad 4.23$$

$$[\hat{R} - \phi^{-1}(1 - \frac{\alpha}{2}) \cdot \sqrt{E\hat{Q}M(\hat{R})}; \hat{R} + \phi^{-1}(1 - \frac{\alpha}{2}) \cdot \sqrt{E\hat{Q}M(\hat{R})}] \quad 4.24$$

Através da experiência do atuário responsável, este pode optar, numa perspetiva mais cautelosa, considerar o provisionamento de um limite superior dos seus respetivos intervalos de confiança com valor elevado.

A maior vantagem em utilizar este método é poder avaliar se a estimativa pontual obtida pelo método de *Chain Ladder* varia, conferindo um grau de confiança às estimativas anteriormente exibidas.

5. CÁLCULO DOS FLUXOS DE CAIXA E O DESDOBRAMENTO DO FATOR CAUDA

Com o reporte em ambiente da Solvência II, para este trabalho são necessários o cálculo dos fluxos de caixa para o desdobramento do fator cauda. Para tal, após o cálculo das provisões com o método de Thomas Mack, com o suporte de Excel serão calculados os fluxos de caixa dos triângulos para estender o triângulo *run-off* e extrapolar um bom fator cauda da matriz.

O fator cauda é o valor representativo dos montantes que serão pagos após o último ano tendo em conta o ano de desenvolvimento que é conhecido ou o último na matriz de triângulo *run-off*.

Durante a análise de uma matriz, se uma seguradora apenas utilizar a informação disponível das matrizes de triângulo *run-off*, estará a assumir que não existem mais pagamentos a serem fechados nos sinistros que se encontravam na linha I e que serão apenas reportados ou liquidados no momento $I+1$. Desta forma, em algumas linhas de negócio, as estimativas calculadas podem encontrar-se menos ajustadas à realidade da seguradora. Assim, o fator cauda irá proporcionar à seguradora as evoluções futuras desconhecidas até agora.

Utilizou-se o *package* ChainLadder no R Cran para calcular as provisões para sinistros e os respetivos intervalos de confiança (Carrato A., 2019). Os fluxos de caixa dos triângulos do montante de indemnização dos sinistros foram calculados, recorrendo ao Excel e ao R Cran em quatro casos (C1, C2, C3 e C4).

No primeiro caso (C1) é considerado um caso *standard* do triângulo sem a hipótese de existir uma cauda nos anos futuros de pagamento de sinistros. Desta forma, é suposto que o ano de desenvolvimento I seja o último ano de pagamentos de sinistros.

No segundo caso (C2) é tida em conta a informação dada pela empresa Seguradora com os valores de provisão caso a caso na extrapolação da cauda. São considerados os dados do primeiro ano de sinistro para o cálculo dos fatores de desenvolvimento que irão extrapolar o resto do triângulo.

No terceiro caso (C3), é utilizado o método de ajustamento logarítmico linear com os coeficientes calculados com o método de *Chain Ladder*. O objetivo é extrapolar 100 períodos (neste caso anos), desde o i -ésimo ano até ao final dos fatores de desenvolvimento. Também se recorreu às ferramentas estatísticas disponíveis no R Cran para realizar o ajustamento logarítmico linear.

Para o último caso (C4) foi utilizado o método apresentado por Neuhaus, W. (2008), que se baseia na extrapolação da cauda através do cálculo de novos fatores de desenvolvimento a partir do ano de ocorrência fixo (escolhido pelo atuário).

Considera-se I o último ano com registo de pagamentos de indemnizações. \hat{g}_k são os novos coeficientes de desenvolvimento.

$\hat{g}_k = f_k$, $k < t < I$, sendo t o ano de desenvolvimento fixado pelo atuário.

$$\hat{g}_k = 1 + \delta(\hat{g}_{k-1} - 1), \text{ para } k > t \text{ com } 0 < \delta < 1 \quad 5.1$$

O valor de δ é obtido de acordo com os dados disponíveis. Devem considerar-se vários cenários para t e para δ , tendo em atenção a experiência do atuário.

As estimativas das provisões para sinistro são, agora, determinadas de acordo com a fórmula do método original.

Se o objetivo for uma cauda curta, δ será mais próximo de 0. Por outro lado, se a linha de negócio em análise tem uma cauda, de um desenvolvimento mais longo, o δ será próximo de 1.

De seguida, utilizar-se-á o *package* de ChainLadder no R Cran para obter as reservas.

6. APLICAÇÃO DO MODELO *THOMAS MACK*

Considere a seguinte matriz de desenvolvimento dos montantes pagos incrementais (como na Tabela 3.1, de sinistros entre 2003 e 2018 (15 anos). A moeda utilizada é Euros e é indicado em milhares de euros.

Tabela 6.1 - Montantes Incrementais Pagos (em milhares de euros)

Período de Ocorrência	Período de Desenvolvimento															
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2003	2332	3051	2382	2121	1235	1055	1302	1149	846	283	166	182	7	27	55	20
2004	1226	3094	2384	2215	1491	2010	746	1626	883	239	34	109	6	221	161	
2005	1399	3378	2620	2081	1193	1645	1010	511	1301	659	591	18	2	11		
2006	1907	3483	1943	1574	760	1434	1510	1011	448	119	1001	120	0			
2007	1701	3378	2200	2250	901	1201	1048	698	683	270	1722	3				
2008	1910	3483	3304	2611	1457	373	1342	1580	258	128	43					
2009	1280	3787	1900	1766	611	1005	380	362	1243	99						
2010	1879	3849	2097	2096	1919	2377	2541	682	60							
2011	1602	3640	2486	1055	987	792	929	784								
2012	1268	2394	1815	748	1882	1982	809									
2013	1169	3680	2240	955	1317	1544										
2014	1670	4118	2028	2903	1357											
2015	1549	2612	1594	1186												
2016	1188	2434	1731													
2017	819	2299														
2018	877															

Para aplicar este método, é necessário em primeiro lugar calcular a matriz de desenvolvimento com os valores acumulados como na Equação 3.2.

Tabela 6.2 - Montantes Pagos Acumulados (em milhares de euros)

Período de Ocorrência	Período de Desenvolvimento															
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2003	2332	5382	7764	9885	11120	12175	13477	14626	15473	15756	15922	16104	16111	16139	16194	16214
2004	1226	4321	6704	8919	10410	12420	13166	14793	15676	15914	15948	16057	16063	16284	16446	
2005	1399	4776	7397	9478	10671	12316	13326	13837	15138	15797	16387	16405	16407	16418		
2006	1907	5390	7332	8906	9666	11100	12610	13621	14069	14187	15188	15308	15308			
2007	1701	5488	7689	9939	10840	12041	13089	13787	14470	14739	16462	16465				
2008	1910	5959	9262	11873	13330	13703	15045	16626	16883	17011	17054					
2009	1280	4729	6629	8395	9006	10012	10392	10754	11997	12096						
2010	1879	5729	7826	9921	11840	14217	16758	17440	17501							
2011	1602	5242	7728	8783	9770	10563	11491	12275								
2012	1268	3662	5477	6225	8107	10089	10899									
2013	1169	4849	7089	8044	9362	10906										
2014	1670	5788	7816	10719	12077											
2015	1549	4161	5755	6941												
2016	1188	3622	5353													
2017	819	3118														
2018	877															

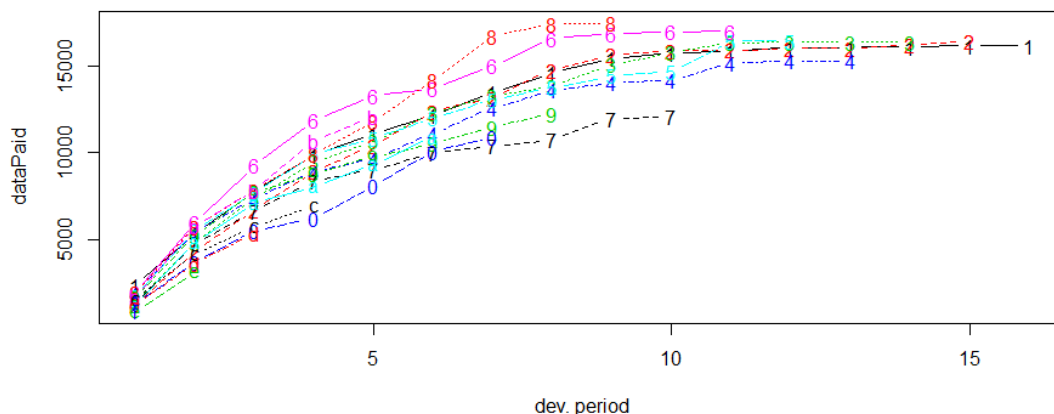


Figura 6.1 - Montantes Pagos Acumulados (*R Cran*)

O gráfico acima apresenta a evolução ao longo dos anos de desenvolvimento dos montantes acumulados das indemnizações. Pode observar-se que as diversas linhas, correspondendo cada uma delas a um ano de sinistro, apresentam um comportamento semelhante.

Os fatores de desenvolvimento do método *Chain Ladder* são calculados aplicando a Equação 3.3 e apresentando abaixo:

Tabela 6.3 - Fatores de Desenvolvimento

k	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
\hat{f}_k	3.154	1.445	1.249	1.14	1.135	1.098	1.07	1.05	1.017	1.038	1.005	1	1.005	1.007	1.001

Verifica-se de seguida os pressupostos para aplicação de *Chain Ladder*.

Nesse sentido, a primeira análise foca-se na adequação dos fatores de desenvolvimento estimados. Desta forma, através da observação dos gráficos abaixo, é possível perceber a relação linear subjacente à existência de proporcionalidade entre os anos de desenvolvimento.

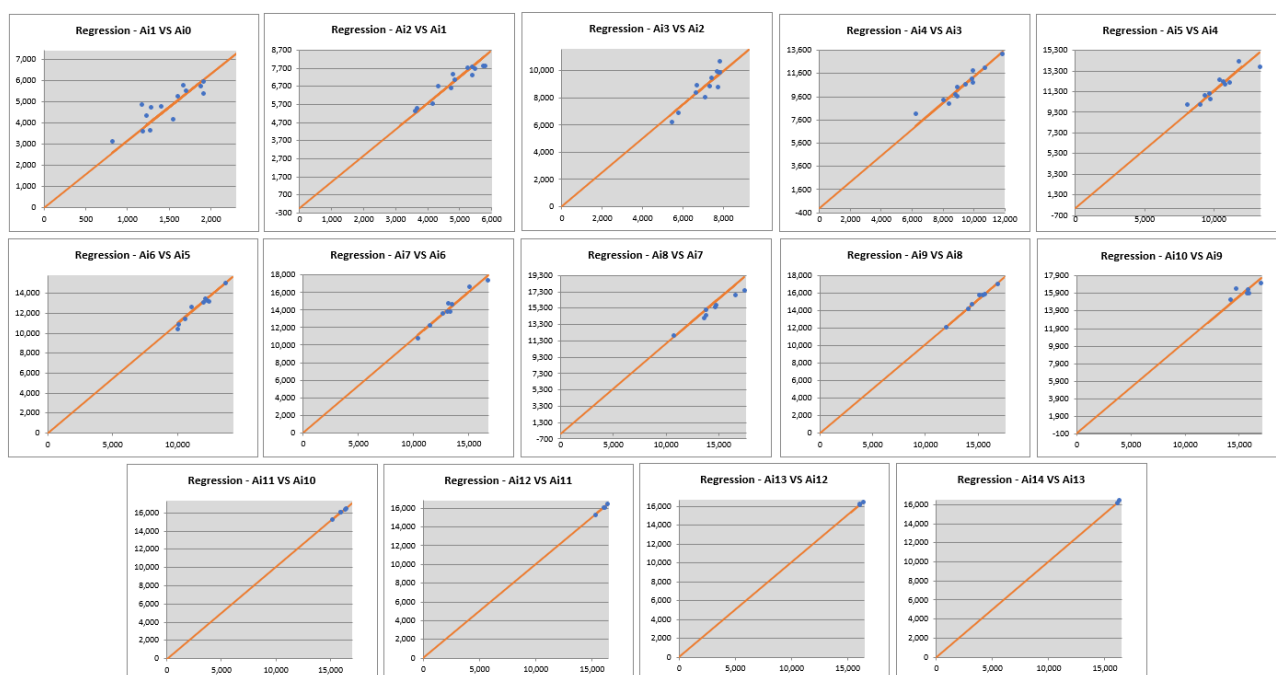


Figura 6.2 - Ajustamento dos dados aos fatores de desenvolvimento

Como pode ser observado, os gráficos acima mostram os dados ordenados $(A_{i,k}, A_{i,k+1})$, permitindo a conclusão de que as retas de declive \hat{f}_k que passam na origem estão bem ajustadas aos dados observados, não havendo desvios significativos. Desta forma, a hipótese da existência de proporcionalidade entre os anos de desenvolvimento não é rejeitada.

Além disto, será ainda necessário verificar a hipótese da não correlação entre os fatores de desenvolvimento individuais, nomeadamente os fatores de desenvolvimento $A_{i,k+1}/A_{i,k}$, como na Equação 4.1. Para tal, o teste de *Spearman* é utilizado.

Tabela 6.4 - Coeficientes dos fatores de desenvolvimento Individuais

Período de Ocorrência	Período de Desenvolvimento															
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2003	2.308	1.443	1.273	1.125	1.095	1.107	1.085	1.058	1.018	1.011	1.011	1.000	1.002	1.003	1.001	
2004	3.524	1.551	1.330	1.167	1.193	1.060	1.124	1.060	1.015	1.002	1.007	1.000	1.014	1.010		
2005	3.414	1.549	1.281	1.126	1.154	1.082	1.038	1.094	1.044	1.037	1.001	1.000	1.001			
2006	2.826	1.360	1.215	1.085	1.148	1.136	1.080	1.033	1.008	1.071	1.008	1.000				
2007	3.226	1.401	1.293	1.091	1.111	1.087	1.053	1.050	1.019	1.117	1.000					
2008	3.120	1.554	1.282	1.123	1.028	1.098	1.105	1.015	1.008	1.003						
2009	3.695	1.402	1.266	1.073	1.112	1.038	1.035	1.116	1.008							
2010	3.049	1.366	1.268	1.193	1.201	1.179	1.041	1.003								
2011	3.272	1.474	1.137	1.112	1.081	1.088	1.068									
2012	2.888	1.496	1.137	1.302	1.244	1.080										
2013	4.148	1.462	1.135	1.164	1.165											
2014	3.466	1.350	1.371	1.127												
2015	2.686	1.383	1.206													
2016	3.049	1.478														
2017	3.807															
2018																

De seguida, o conjunto de dados é testado para a inexistência de correlações entre os fatores de desenvolvimento individuais adjacentes. Para realizar o teste de *Spearman*, as matrizes seguintes foram construídas com os números de ordem $(r_{i,k})$ e $(s_{i,k})$ com base nos fatores de desenvolvimento individuais das figuras abaixo.

Tabela 6.5 - (A): Matriz $r_{i,k}$ e (B): Matriz $s_{i,k}$

Período de Ocorrência	Período de Desenvolvimento															
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2003	7	8	6	3	8	7	5	3	5	3	2	1	1			
2004	13	12	10	9	2	9	6	4	1	3	2	3	2			
2005	12	9	7	7	4	2	7	7	4	2	1	1				
2006	2	5	2	6	9	6	3	3	5	4	1					
2007	5	11	3	4	5	4	4	6	6	1						
2008	14	10	5	1	7	8	2	1	2							
2009	6	6	1	5	1	1	8	2								
2010	3	7	11	10	10	3	1									
2011	9	2	4	2	6	5										
2012	11	3	12	11	3											
2013	8	1	9	8												
2014	1	13	8													
2015	4	4														
2016	10															
2017																
2018																

(A)

Período de Ocorrência	Período de Desenvolvimento															
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2003	1	7	7	6	3	3	4	4	4	2	4	3	1	1		
2004	12	12	11	9	8	8	8	5	3	1	2	2	2			
2005	10	11	8	7	7	7	2	6	6	3	1	1				
2006	3	2	4	2	6	2	5	2	2	4	3					
2007	8	5	10	3	4	6	4	3	5	5						
2008	7	13	9	5	1	4	7	1	1							
2009	13	6	5	1	5	9	1	7								
2010	5	3	6	10	9	1	3									
2011	9	9	2	4	2	5										
2012	4	10	3	11	10											
2013	14	8	1	8												
2014	11	1	12													
2015	2	4														
2016	6															
2017																
2018																

(B)

Os resultados foram obtidos na tabela abaixo, com a Equação 4.2:

Tabela 6.6 - Coeficientes de Correlação de Spearman

I	15														
k	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
T(k)		0.213	-0.006	0.027	0.718	-0.203	-0.217	-0.333	0.321	0.429	-0.500	0.500	0.500	1.000	
V(T(k))		0.077	0.083	0.091	0.100	0.111	0.125	0.143	0.167	0.200	0.25	0.333	0.500	1.000	

T é obtido em resultado da Equação 4.4, sendo $T = 0.10832$.

Na ausência de hipótese de correlações espera-se que $E(T) = 0$ e $V(T) = 0.01098$ como na Equação 4.3.

Com um intervalo de confiança de 95% a estimativa T será

$[-0.070705738; 0.070705738]$ como na Equação 4.5.

Como é possível observar, a estimativa T não se encontra dentro do intervalo, por isso a hipótese da não correlação entre os fatores de desenvolvimento é rejeitada, por um nível de significância de 5%.

De seguida testa-se a independência entre os diferentes anos de ocorrência (as linhas da matriz acumulada). Para esta hipótese, é feita uma análise, coluna a coluna de todos os fatores de desenvolvimento individuais para identificar quais destes pertencem aos conjuntos P e G. Os fatores indicados com asterisco (*) representam os elementos desprezados quando os fatores de desenvolvimento têm um número ímpar, ou seja, o valor mediano.

Os resultados foram obtidos abaixo:

Tabela 6.7 - Conjuntos P e G

Período de Ocorrência	Período de Desenvolvimento															
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2003		P	P	G	P	P	G	G	G	G	P	G	G	*	P	*
2004		G	G	G	G	G	P	G	G	*	P	*	G	G	G	
2005		G	G	G	G	G	P	P	G	G	G	P	P	P		
2006		P	P	P	P	*	G	G	P	P	G	G	P			
2007		*	P	G	P	P	P	P	P	G	G	P				
2008		P	G	G	P	P	G	G	P	P	P					
2009		G	P	P	P	P	P	P	G	G						
2010		P	P	*	G	G	G	P	P							
2011		G	G	P	P	P	G	*								
2012		P	G	P	G	G	P									
2013		G	G	P	G	G										
2014		G	P	G	G											
2015		P	P	P												
2016		P	G													
2017		G														
2018																

Seendo A_k $0 \leq k \leq 15$ a diagonal do triângulo acima, P_k será o número de elementos P que pertencem a A_k e G_k será o número de elementos G que pertencem a A_k . Desta forma, tem-se:

Tabela 6.8 - Momentos das Z_k e Z

k	P_k	G_k	Z_k	n	m	$E(Z_k)$	$V(Z_k)$	$(P_k+G_k)/2$
0	1	0	0	1	0	0.00	0.00	0.5
1	1	1	1	2	0	0.50	0.25	1
2	0	3	0	3	1	0.75	0.19	1.5
3	2	2	2	4	1	1.25	0.44	2
4	2	2	2	4	1	1.25	0.44	2
5	3	3	3	6	2	2.06	0.62	3
6	2	5	2	7	3	2.41	0.55	3.5
7	4	3	3	7	3	2.41	0.55	3.5
8	5	3	3	8	3	2.91	0.80	4
9	5	3	3	8	3	2.91	0.80	4
10	5	6	5	11	5	4.15	0.92	5.5
11	5	6	5	11	5	4.15	0.92	5.5
12	7	5	5	12	5	4.65	1.17	6
13	6	8	6	14	6	5.53	1.35	7
14	7	6	6	13	6	5.03	1.10	6.50
Total			30			31.73	7.61	

O valor $E(Z_k) = 31.73$ como na Equação 4.7 e o valor da $V(Z_k) = 7.61$ como na Equação 4.8.

O Intervalo de Confiança de 95% será [26.3170;37.1342], com o cálculo da Equação 4.6.

A estimativa de Z é $Z = 30$, pertencendo ao intervalo de confiança.

Desta forma, a hipótese não é rejeitada.

A última hipótese para a verificação do modelo de *Thomas Mack* serve para garantir que os estimadores centrais obtidos têm a menor variância (Equação 4.10).

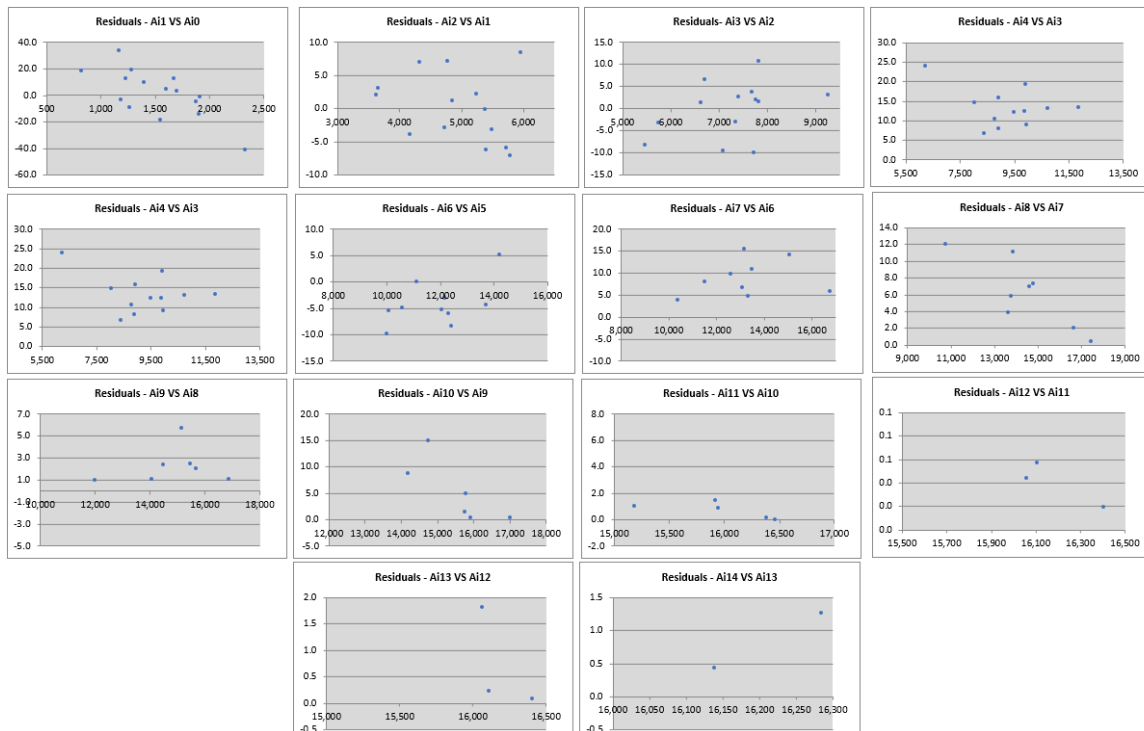


Figura 6.3 - Resíduos Ponderados

Os gráficos apresentam as ponderações residuais que permitem apurar se os estimadores de fatores de desenvolvimento alcançados são os que originam uma menor variância.

Observando todos os gráficos da Figura 6.3, a dispersão das ponderações residuais não mostra uma tendência, uma vez que os dados estão dispersos aleatoriamente.

Por isso, a hipótese não se rejeita.

Com os resultados, pode afirmar-se que as hipóteses implícitas no modelo de *Thomas Mack* encontram-se satisfatoriamente verificadas no conjunto de dados, apesar da condição de não correlação entre os fatores de desenvolvimento não se verificar.

Tendo em atenção a ilustração dos objetivos apresentados neste trabalho, irá dar-se continuidade ao processo. Noutra situação seria necessário aplicar outro modelo aos dados, pois não está verificado um dos pressupostos do modelo.

7. RESULTADOS E DISCUSSÃO

O impacto do *ultimate* será avaliado na obtenção das provisões para sinistros. Para tal, irão aplicar-se os quatros casos (C1, C2, C3 e C4) apresentados na secção 5 aos dados já apresentados na secção anterior.

Para a presente dissertação, foram utilizados os pagamentos dos sinistros referentes à realidade de uma Seguradora, desde o ano de 2003 até 2018 (15 anos) com os danos corporais na Linha de Negócio Automóvel. Para garantir a confidencialidade da mesma, foi aplicada uma percentagem sobre os dados originais.

7.1. CASO 1

Considerando que os pagamentos de sinistro se resolvem em 15 anos, através da Tabela 7.1 pode observar-se que ao fim de 6 anos é que estão efetuadas cerca de 75% dos pagamentos. Utiliza-se a estimativa dos fatores de desenvolvimento (f_k), com os montantes pagos acumulados $A_{i,k+1}/A_{i,k}$ da Equação 3.2 já calculados na secção anterior (Tabela 6.3).

Tabela 7.1 - Lag Factor (%)

k	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Lag Factor (%)	10%	32%	47%	58%	66%	75%	83%	89%	93%	95%	98%	99%	99%	99%	100%

Através do *Lag Factor* é possível analisar que também existe uma tendência para um aumento gradual do pagamento dos sinistros ocorridos ao longo dos anos. É ainda na passagem do ano de desenvolvimento 0 para o 1 que se verifica uma maior acumulação de pagamentos efetuados até ao primeiro ano (10% para 32%). A partir do ano de desenvolvimento 10 e até ao fim, a acumulação continua a aumentar, mas em menor proporção, mantendo-se igual até ao final dos anos de desenvolvimento.

Depois dos cálculos relacionados com os fatores de desenvolvimento, procede-se à projeção do valor dos sinistros a ser pago no futuro (o destaque azul do triângulo inferior da matriz).

Tabela 7.2 - Matriz Incremental preenchida (C1)

Período de Ocorrência	Período de Desenvolvimento															
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2003	2332	3051	2382	2121	1235	1055	1302	1149	846	283	166	182	7	27	55	20
2004	1226	3094	2384	2215	1491	2010	746	1626	883	239	34	109	6	221	161	20
2005	1399	3378	2620	2081	1193	1645	1010	511	1301	659	591	18	2	11	110	20
2006	1907	3483	1943	1574	760	1434	1510	1011	448	119	1001	120	0	82	103	19
2007	1701	3787	2200	2250	901	1201	1048	698	683	270	1722	3	4	88	110	20
2008	1910	4049	3304	2611	1457	373	1342	1580	258	128	43	92	4	92	115	21
2009	1280	3449	1900	1766	611	1005	380	362	1243	99	460	3	3	67	84	16
2010	1879	3849	2097	2096	1919	2377	2541	682	60	666	692	102	5	101	127	24
2011	1602	3640	2486	1055	987	792	929	784	608	223	499	74	3	73	92	17
2012	1268	2394	1815	748	1882	1982	809	767	578	212	474	70	3	69	87	16
2013	1169	3680	2240	955	1317	1544	1068	843	635	233	521	77	3	76	96	18
2014	1670	4118	2028	2903	1357	1632	1342	1060	798	293	655	97	4	96	120	22
2015	1549	2612	1594	1186	944	1065	876	692	521	191	428	63	3	63	79	15
2016	1188	2434	1731	1335	910	1027	845	667	502	184	412	61	3	60	76	14
2017	819	2299	1386	1123	766	864	711	561	423	155	347	51	2	51	64	12
2018	877	1889	1230	997	679	766	631	498	375	138	308	45	2	45	57	10

De seguida, os fluxos de Caixa são calculados somando a diagonal da matriz, como apresentado abaixo:

Tabela 7.3 - Fluxos de Caixa (C1)

k	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033
Fluxo de Caixa	10984	11064	8225	6397	4866	3759	2836	1817	1190	756	562	209	143	123	68	10

De acordo com os resultados apresentados, a partir deste método pode verificar-se que 2033 será o último ano em que a seguradora terá encargos com sinistros de 2018 para pagar.

Os resultados do método de *Thomas Mack* foram estimados com a utilização do *Package ChainLadder* no R Cran, onde foram estimadas as reservas e os limites inferiores e superiores com um intervalo de confiança de 95%.

Tabela 7.4 - Aplicação do método de Chain Ladder (C1)

	Reservas	Erro Padrão	Limite Inferior	Limite Superior
Total	41 706	14 344	13 592	69 820

Na análise dos valores obtidos, considerando que os pagamentos se resolvem em 15 anos, estima-se que para o final do exercício de 2018, será necessário um valor de reserva de aproximadamente 41 706 milhares de euros para regularizar todos os sinistros ocorridos entre 2003 e 2018, mas ainda pendentes.

O Erro Quadrático Médio associado à provisão estimada é de 14 344 milhares de euros, representando 34% das reservas. O intervalo de confiança de 95% é de [13 592; 69 820].

7.2. CASO 2

Considerando a informação dada pela empresa Seguradora com os valores de provisão, foram adicionados os dados do primeiro ano do sinistro para extrapolar o fator cauda para o máximo do resto do triângulo.

O cálculo dos fatores de desenvolvimento foi feito em R Cran através da função *MackChainLadder* do *package* ChainLadder.

Depois do cálculo dos novos fatores de desenvolvimento, procede-se à projeção do valor dos sinistros a ser pago no futuro (o destaque azul do triângulo inferior da matriz) e respetivos fluxos de caixa.

Tabela 7.5 - Novos Fatores de Desenvolvimento (C2)

k	0	1	2	3	4	5	6	7	8	9	10	11	12
\hat{f}_k	3.53717	1.444621	1.249409	1.136025	1.135121	1.097925	1.070395	1.049557	1.017313	1.038069	1.005408	1.000239	1.005343
	13	14	15										
	1.00667	1.001229	1.053781										

Como não existe informação adicional sobre a expectativa de pagamentos futuros, a provisão futura é calculada com base no fator de desenvolvimento 1.053781 (ano 15). Desta forma, a coluna >16 é criada. Assim, fazendo a diferença entre as colunas >16 e 15, obtém-se a provisão futura incremental (coluna ≥16) para o cálculo dos fluxos de caixa.

Ao ser considerada, a provisão futura é decomposta manualmente para os 10 anos de desenvolvimento seguintes, como é possível observar-se na tabela abaixo. Pela escolha do atuário, no ano de desenvolvimento 16 iriam existir para todos os anos de ocorrência 100 milhares de euros a pagar em compensações aos clientes.

Desse período para a frente os resultados seguintes foram estimados com a factorização de uma constante para cada ano de sinistros, sendo estes valores próximos de 1.

Tabela 7.6 - Cálculo de Provisão Futura (C2)

Período de Ocorrência	Período de Desenvolvimento																	
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	>16	≥16
2003	2332	5382	7764	9885	11120	12175	13477	14626	15473	15756	15922	16104	16111	16139	16194	16214	17086	872
2004	1226	4321	6704	8919	10410	12420	13166	14793	15676	15914	15948	16057	16063	16284	16446	16466	17352	886
2005	1399	4776	7397	9478	10671	12316	13326	13837	15138	15797	16387	16405	16407	16418	16528	16548	17438	890
2006	1907	5390	7332	8906	9666	11100	12610	13621	14069	14187	15188	15308	15308	15390	15492	15512	16346	834
2007	1701	5488	7689	9939	10840	12041	13089	13787	14470	14739	16462	16465	16469	16557	16667	16688	17585	897
2008	1910	5959	9262	11873	13330	13703	15045	16626	16883	17011	17054	17146	17150	17242	17357	17378	18313	935
2009	1280	4729	6629	8395	9006	10012	10392	10754	11997	12096	12556	12624	12627	12695	12780	12795	13483	688
2010	1879	5729	7826	9921	11840	14217	16758	17440	17501	17804	18482	18582	18586	18685	18810	18833	19846	1013
2011	1602	5242	7728	8783	9770	10563	11491	12275	12883	13106	13605	13679	13682	13755	13847	13864	14610	746
2012	1268	3662	5477	6225	8107	10089	10899	11666	12244	12456	12931	13000	13004	13073	13160	13176	13885	709
2013	1169	4849	7089	8044	9362	10906	11974	12817	13452	13685	14206	14283	14286	14362	14458	14476	15255	779
2014	1670	5788	7816	10719	12077	13709	15051	16111	16909	17202	17857	17953	17958	18054	18174	18196	19175	979
2015	1549	4161	5755	6941	7885	8951	9827	10519	11040	11231	11659	11722	11725	11787	11866	11881	12520	639
2016	1188	3622	5353	6688	7598	8624	9469	10136	10638	10822	11234	11295	11297	11358	11434	11448	12063	616
2017	819	3118	4504	5628	6393	7257	7968	8529	8951	9106	9453	9504	9506	9557	9621	9633	10151	518
2018	877	3102	4481	5599	6361	7220	7927	8485	8906	9060	9405	9456	9458	9508	9572	9584	10099	515

Assim, com a decomposição destes valores os fluxos de caixa estendem-se durante mais tempo, sendo que a partir do ano de desenvolvimento 16 estes vão baixando até chegarem ao valor provisionado.

Tabela 7.7 - Matriz Incremental Preenchida (C2)

Período de Ocorrência	Período de Desenvolvimento																									
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
2003	2332	3050	2382	2121	1235	1055	1302	1149	847	283	166	182	7	28	55	20	100	97	94	91	88	85	83	80	78	75
2004	1226	3095	2383	2215	1491	2010	746	1627	883	238	34	109	6	221	162	20	100	97	95	92	90	87	85	82	80	78
2005	1399	3377	2621	2081	1193	1645	1010	511	1301	659	590	18	2	11	110	20	100	97	95	92	90	88	85	83	81	79
2006	1907	3483	1942	1574	760	1434	1510	1011	448	118	1001	120	0	82	103	19	100	96	92	88	85	81	78	75	71	69
2007	1701	3787	2201	2250	901	1201	1048	698	683	269	1723	3	4	88	110	20	100	98	95	93	91	88	86	84	82	80
2008	1910	4049	3303	2611	1457	373	1342	1581	257	128	43	92	4	92	115	21	100	98	97	96	94	93	91	90	89	87
2009	1280	3449	1900	1766	611	1006	380	362	1243	99	460	68	3	67	85	16	100	91	83	76	70	64	58	53	48	44
2010	1879	3850	2097	2095	1919	2377	2541	682	61	303	678	100	4	99	125	23	100	91	83	75	68	62	56	51	47	42
2011	1602	3640	2486	1055	987	793	928	784	608	223	499	74	3	73	92	17	100	93	87	81	76	71	66	61	57	53
2012	1268	2394	1815	748	1882	1982	810	767	578	212	474	70	3	69	87	16	100	92	85	78	72	66	61	56	52	48
2013	1169	3680	2240	955	1318	1544	1068	843	635	233	521	77	3	76	96	18	100	94	89	84	79	75	70	66	63	59
2014	1670	4118	2028	2903	1358	1632	1342	1060	798	293	655	97	4	96	120	22	100	100	99	99	98	98	97	97	96	96
2015	1549	2612	1594	1186	944	1065	876	692	521	191	428	63	3	63	79	15	100	90	80	72	64	58	51	46	41	37
2016	1188	2434	1731	1335	910	1027	845	667	502	184	412	61	3	60	76	14	100	89	79	70	62	55	48	43	38	34
2017	819	2299	1386	1123	766	864	711	561	423	155	347	51	2	51	64	12	100	84	71	60	50	42	35	30	25	21
2018	877	2225	1379	1118	762	859	707	558	420	154	345	51	2	51	63	12	100	84	70	59	50	42	35	29	25	21

De acordo com os resultados apresentados, a partir deste método pode verificar-se que 2043 será o último ano em que a seguradora terá encargos com sinistros de 2018 para pagar.

Tabela 7.8 - Fluxo de Caixa (C2)

k	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030
Fluxo de Caixa	10988	11137	8622	6807	5331	4321	3467	2519	1954	1559	1451	1051	961

2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043
931	868	777	720	612	518	422	342	257	140	88	46	21

Na análise dos valores obtidos com a utilização do *Package ChainLadder* no R Cran, considerando a informação da seguradora com os valores de provisão, estima-se que para o final do exercício de 2018, será necessário um valor de reserva de aproximadamente 54 166 milhares de euros para regularizar todos os sinistros ocorridos entre 2003 e 2018, mas ainda pendentes.

7.3. CASO 3

Foi considerado um ajustamento logarítmico linear nos coeficientes do método de *Chain Ladder*. Para o ajustamento, são calculados novos fatores de desenvolvimento, estando os seus resultados na tabela abaixo. Após o cálculo, foi feita uma regressão logarítmica nos fatores de idade para idade e os gráficos abaixo identificam a extrapolação e o padrão desenvolvimento que é esperado dos sinistros.



Figura 7.1- Extrapolação Log-Linear dos fatores de idade para idade (C1)

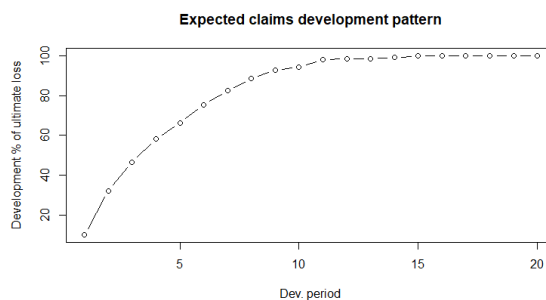


Figura 7.2 - Padrão do Desenvolvimento Esperado dos Sinistros (C3)

Tabela 7.9 - Novos Fatores de Desenvolvimento (C3)

k	0	1	2	3	4	5	6	7	8	9	10
\hat{f}_k	3.53717	1.444621	1.249409	1.136025	1.135121	1.097925	1.070395	1.049557	1.017313	1.038069	1.005408

11	12	13	14	15	16	17	18	19	20	21
1.000239	1.005343	1.006673	1.001229	1.00069	1.000428	1.000264	1.000163	1.000101	1.000062	1.000038

22	23	24	25	26	27
1.000024	1.000015	1.000009	1.000006	1.000003	1.000002

Após calcular os fatores de desenvolvimento, procedeu-se à projeção do valor dos sinistros a ser pago no futuro (o destaque azul do triângulo inferior da matriz) e respetivos fluxos de caixa, utilizando os novos fatores de desenvolvimento calculados até ao ano de desenvolvimento 28.

Tabela 7.10 - Matriz Incremental Preenchida (C3)

Período de Ocorrência	Período de Desenvolvimento																													
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	
2003	2332	3050	2382	2121	1235	1055	1302	1149	847	283	166	182	7	28	55	20	11	7	4	3	2	1	1	0	0	0	0	0	0	0
2004	1226	3095	2383	2215	1491	2010	746	1627	883	238	34	109	6	221	162	20	11	7	4	3	2	1	1	0	0	0	0	0	0	0
2005	1399	3377	2621	2081	1193	1645	1010	511	1301	659	590	18	2	11	110	20	11	7	4	3	2	1	1	0	0	0	0	0	0	0
2006	1907	3483	1942	1574	760	1434	1510	1011	448	118	1001	120	0	82	103	19	11	7	4	3	2	1	1	0	0	0	0	0	0	0
2007	1701	3787	2201	2250	901	1201	1048	698	683	269	1723	3	4	88	110	20	12	7	4	3	2	1	1	0	0	0	0	0	0	0
2008	1910	4049	3303	2611	1457	373	1342	1581	257	128	43	92	4	92	115	21	12	7	5	3	2	1	1	0	0	0	0	0	0	0
2009	1280	3449	1900	1766	611	1006	380	362	1243	99	460	68	3	67	85	16	9	5	3	2	1	1	0	0	0	0	0	0	0	0
2010	1879	3850	2097	2095	1919	2377	2541	682	61	303	678	100	4	99	125	23	13	8	5	3	2	1	1	0	0	0	0	0	0	0
2011	1602	3640	2486	1055	987	793	928	784	608	223	499	74	3	73	92	17	10	6	4	2	1	1	1	0	0	0	0	0	0	0
2012	1268	2394	1815	748	1882	1982	810	767	578	212	474	70	3	69	87	16	9	6	3	2	1	1	1	0	0	0	0	0	0	0
2013	1169	3680	2240	955	1318	1544	1068	843	635	233	521	77	3	76	96	18	10	6	4	2	1	1	1	0	0	0	0	0	0	0
2014	1670	4118	2028	2903	1358	1632	1342	1060	798	293	655	97	4	96	120	22	13	8	5	3	2	1	1	0	0	0	0	0	0	0
2015	1549	2612	1594	1186	944	1065	876	692	521	191	428	63	3	63	79	15	8	5	3	2	1	1	0	0	0	0	0	0	0	0
2016	1188	2434	1731	1335	910	1027	845	667	502	184	412	61	3	60	76	14	8	5	3	2	1	1	0	0	0	0	0	0	0	0
2017	819	2299	1386	1123	766	864	711	561	423	155	347	51	2	51	64	12	7	4	3	2	1	1	0	0	0	0	0	0	0	0
2018	877	2225	1379	1118	762	859	707	558	420	154	345	51	2	51	63	12	7	4	3	2	1	1	0	0	0	0	0	0	0	0

De acordo com os resultados apresentados, a partir deste método pode verificar-se que 2046 será o último ano em que a seguradora terá encargos com sinistros de 2018 para pagar.

Tabela 7.11 - Fluxos de Caixa (C3)

k	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032
Fluxos de Caixa	10988	11048	8444	6539	4973	3877	2938	1903	1265	800	626	241	172	154	99

2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046
33	20	12	8	5	3	2	1	1	0	0	0	0	0

Na análise dos valores obtidos com a utilização do *Package* ChainLadder no R Cran, considerando o ajustamento logarítmico linear, estima-se que para o final do exercício de 2018, será necessário um valor de reserva de aproximadamente 42 096 milhares de euros para regularizar todos os sinistros ocorridos entre 2003 e 2018, mas ainda pendentes.

7.4. CASO 4

Considerando o modelo apresentado por Neuhaus, W. (2008) (Equação 5.1) foram utilizados três tipos de ponderações (δ) para testar qual deles seria o mais adequado para extrapolar a cauda. Os valores utilizados foram 0.5, 0.7 e 0.8.

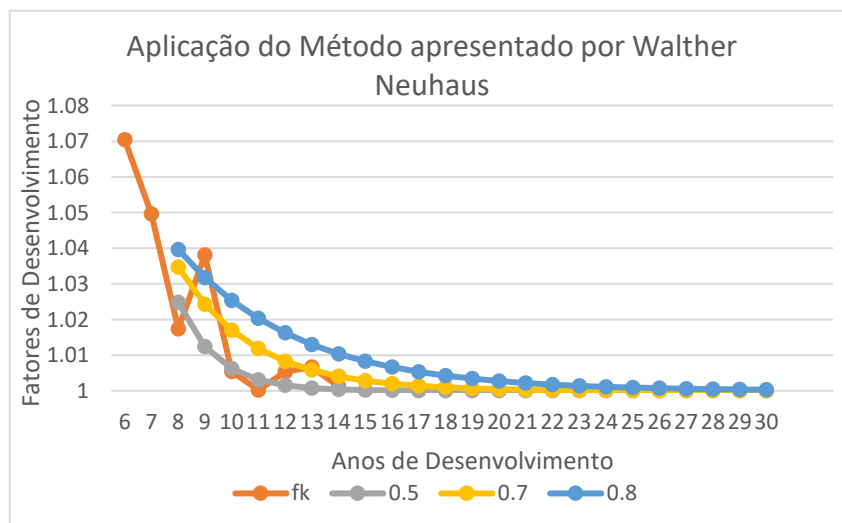


Figura 7.3 - Novos Fatores de Desenvolvimento (C4)

Com base na experiência do atuário, a melhor aplicação é a escolha da ponderação 0.7 (linha amarela). Esta análise é qualitativa com base na observação do gráfico e na expectativa da evolução deste ramo em particular. A ponderação 0.5 (linha cinzenta) não é a mais adequada para extrapolar o máximo possível da cauda, visto que se verifica que esta terminaria nos mesmos anos de desenvolvimento que previamente foram dados. Por outro lado, também a ponderação 0.8 (linha azul), não é tão adequada pois afasta-se demasiado dos fatores observados (linha laranja).

Desta forma, obtiveram-se os novos fatores de desenvolvimento dos anos reais da ocorrência, identificados no gráfico da Figura 7.3.

Tabela 7.12 - Novos Fatores de Desenvolvimento (C4)

k	0	1	2	3	4	5	6	7	8	9	10
\hat{f}_k	3.153716	1.444621	1.249409	1.136025	1.135121	1.097925	1.070395	1.049557	1.03469	1.024283	1.016998
11	12	13	14	15	16	17	18	19	20	21	
1.011899	1.008329	1.005830345	1.004081242	1.002856869	1.002	1.0014	1.00098	1.000686	1.00048	1.000336	
22	23	24	25	26	27	28	29	30			
1.000235	1.000165	1.000115	1.000081	1.000056	1.000040	1.000028	1.000019376	1.000013563			

Após calcular os fatores de desenvolvimento, procedeu-se à projeção do valor dos sinistros a ser pago no futuro (o destaque azul do triângulo inferior da matriz) e respetivos fluxos de caixa, utilizando os novos fatores de desenvolvimento que se encontram na Tabela 7.12.

Tabela 7.13 - Matriz Incremental Preenchida (C4)

Período de Ocorrência	Período de Desenvolvimento																																
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	
2003	2332	3050	2382	2121	1235	1055	1302	1149	847	283	166	182	7	28	55	20	46	33	23	16	11	8	5	4	3	2	1	1	1	0	0	0	
2004	1226	3095	2383	2215	1491	2010	746	1627	883	238	34	109	6	221	162	67	47	33	23	16	11	8	6	4	3	2	1	1	1	0	0	0	
2005	1399	3377	2621	2081	1193	1645	1010	511	1301	659	590	18	2	11	96	67	47	33	23	16	11	8	6	4	3	2	1	1	1	0	0	0	
2006	1907	3483	1942	1574	760	1434	1510	1011	448	118	1001	120	0	128	90	63	45	31	22	15	11	8	5	4	3	2	1	1	1	0	0	0	
2007	1701	3787	2201	2250	901	1201	1048	698	683	269	1723	3	196	139	98	69	48	34	24	17	12	8	6	4	3	2	1	1	1	0	0	0	
2008	1910	4049	3303	2611	1457	373	1342	1581	257	128	43	290	206	146	103	73	51	36	25	18	12	9	6	4	3	2	1	1	1	0	0	0	
2009	1280	3449	1900	1766	611	1006	380	362	1243	99	294	211	150	106	75	53	37	26	18	13	9	6	4	3	2	2	1	1	1	0	0	0	
2010	1879	3850	2097	2095	1919	2377	2541	682	61	607	440	315	224	159	112	79	56	39	27	19	13	9	7	5	3	2	2	1	1	1	0	0	
2011	1602	3640	2486	1055	987	793	928	784	608	447	324	232	165	117	83	58	41	29	20	14	10	7	5	3	2	2	1	1	1	0	0	0	
2012	1268	2394	1815	748	1882	1982	810	767	578	425	308	221	157	111	79	55	39	27	19	13	9	7	5	3	2	2	1	1	1	0	0	0	
2013	1169	3680	2240	955	1318	1544	1068	843	635	467	338	242	173	122	86	61	43	30	21	15	10	7	5	4	2	2	1	1	1	0	0	0	
2014	1670	4118	2028	2903	1358	1632	1342	1060	798	587	425	305	217	154	108	76	54	38	26	19	13	9	6	4	3	2	2	1	1	1	0	0	0
2015	1549	2612	1594	1186	944	1065	876	692	521	383	277	199	142	100	71	50	35	25	17	12	8	6	4	3	2	1	1	1	0	0	0	0	
2016	1188	2434	1731	1335	910	1027	845	667	502	369	267	192	136	97	68	48	34	24	17	12	8	6	4	3	2	1	1	1	0	0	0	0	
2017	819	2299	1386	1123	766	864	711	561	423	311	225	161	115	81	57	40	28	20	14	10	7	5	3	2	2	1	1	1	0	0	0	0	
2018	877	1889	1230	997	679	766	630	498	375	275	200	143	102	72	51	36	25	18	12	9	6	4	3	2	1	1	1	1	0	0	0	0	

De acordo com os resultados apresentados, a partir deste método pode verificar-se que 2049 será o último ano em que a seguradora terá encargos com sinistros de 2018 para pagar.

Tabela 7.14 - Fluxos de Caixa (C4)

k	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030
Fluxos de Caixa	10988	11353	8771	6984	5504	4412	3325	2454	1779	1276	909	644	455
2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	
320	225	158	111	78	54	38	26	18	13	9	6	4	
2044	2045	2046	2047	2048	2049								
3	2	1	1	0	0								

Na análise dos valores obtidos com a utilização do *Package ChainLadder* no R Cran, considerando o modelo apresentado por Neuhaus W. (2008) , estima-se que para o final do exercício de 2018, será necessário um valor de reserva de aproximadamente 48 933 milhares de euros para regularizar todos os sinistros ocorridos mas ainda pendentes de 2003 a 2018.

Além do cálculo de reservas destes quatro casos, foram também consideradas as Reservas Disponíveis da Empresa.

Na tabela seguinte pode observar-se os valores das reservas dos quatro casos em estudo e as Reservas Disponíveis da Empresa. Verifica-se que o valor total da reserva estimada que a Companhia precisa constituir para cobrir os passivos por sinistros que têm pagamento ainda pendentes em 2018 é de 44 708 milhares de euros.

Tabela 7.15 - Resultados das Reservas Obtidas

Ano de Ocorrência	Reservas				Empresa
	C1	C2	C3	C4	
2003	0	872	29	155	872
2004	20	906	50	225	13
2005	130	1020	160	321	279
2006	204	1038	232	429	91
2007	223	1120	253	663	14
2008	324	1259	356	989	446
2009	699	1387	723	1012	378
2010	1332	2345	1336	2122	471
2011	1589	2335	1614	2170	3517
2012	2277	2986	2301	2830	1253
2013	3570	4349	3596	4177	1904
2014	6119	7098	6152	6882	5758
2015	4940	5579	4961	5438	5440
2016	6094	6710	6115	6574	4612
2017	6515	7033	6533	6919	8627
2018	7669	8129	7685	8028	11033
Total	41705	54166	42096	48933	44708

É possível observar-se que para o Caso 1 e para o Caso 3 o valor de reservas disponíveis pela Companhia de Seguros é superior à reserva estimada, respetivamente em 3 003 e 2 612 milhares de euros. Estes resultados mostram que, em comparação com o método aplicado, a estrutura obtida está alinhada com o valor obtido pela empresa. Assim, as reservas disponíveis pela empresa são suficientes para atender às suas necessidades. Além disso, o valor total de Reservas Disponíveis da Companhia encontra-se dentro do intervalo de confiança de 95% calculado na Tabela 7.4.

Por outro lado, é possível observar-se que para o Caso 2 e para o Caso 4 a reserva disponível pela empresa é inferior à reserva estimada em 9 458 e 4 225 milhares de euros, respetivamente. Comparando com o método aplicado, as reservas dadas pela empresa mostram-se insuficientes para atender às suas necessidades.

Posto isto, compara-se de seguida os quatro casos entre si, de forma a decidir-se qual o mais adequado para a empresa em estudo aplicar.

O Caso 1 é considerado um caso valor de referência, pois a seguradora tem um provisionamento para os sinistros que ocorreram em 2003 que vai para além dos 15 anos considerados. Contudo como é um caso que não considera o fator cauda, este não deve ser escolhido como o melhor método para a seguradora adequar as suas provisões para sinistros.

Além disso, é possível observar que o valor das reservas dos casos 2, 3 e 4 são superiores aos valores das reservas do Caso 1. Sendo que estes casos consideraram o fator cauda, a decisão do atuário em escolher qualquer um destes métodos é válida.

Começando por analisar as reservas do Caso 2, o montante de reservas referente a sinistros ocorridos em 2003 é elevado comparativamente aos outros casos, mas igual à reserva provisionada pela empresa no mesmo ano (872 milhares de euros).

Contudo, nos anos seguintes, as reservas provisionadas pela empresa são mais baixas que as reservas provisionadas no Caso 2, que geram um montante maior de reservas a um maior ritmo ao longo dos anos de ocorrência, verificando-se um impacto maior acarretando um encargo excessivo para a empresa.

Esta situação deve-se ao facto de o primeiro ano de ocorrência ser um caso atípico, com uma grande provisão. Tendo em conta um coeficiente maior, no Caso 2 houve uma projeção de um elevado valor de provisionamento para o resto dos anos, afetando o triângulo, sendo estes valores maiores que nas reservas provisionadas pela empresa. Desta forma, na perspetiva do atuário, utilizando o Caso 2, a seguradora estaria em grande esforço.

Para o Caso 3, ao ser um ajustamento logarítmico, pode observar-se que a cauda por ser longa demora mais tempo a obter as reservas necessárias para pagar as compensações necessárias aos clientes durante os primeiros anos de ocorrência. Consegue-se observar através do primeiro ano que a reserva é muito mais baixa que a provisão para sinistros da empresa (isto é, uma situação contrária ao que sucede no Caso 2), tornando-se numa situação irregular. A partir do ano de ocorrência de 2009, as reservas do Caso 3 começam a ter valores superiores aos das reservas provisionadas pela empresa. Todavia, pertinente será realçar que ainda assim os valores que este caso apresenta encontram-se abaixo dos que a seguradora está a provisionar.

Tal como no Caso 2, o que ocorre no Caso 3 é considerado algo particular. A empresa ao ter uma provisão no primeiro ano de ocorrência (2003) de 872 milhares de euros significa que nesse ano ocorreu algum sinistro excecionalmente grave. Este tipo de sinistros envolve processos jurídicos em tribunal que fazem com que a empresa tenha de se salvaguardar caso tenham de vir a pagar um valor maior do que aquele que foi calculado.

Assim, a escolha do atuário seria o Caso 4. Apesar de ser um caso em que o total de reservas é superior às reservas provisionadas pela empresa, a opinião do atuário passa por descobrir o coeficiente que se ajusta melhor e dos quatros casos este é aquele que dá resultados mais coerentes. Ao longo dos anos, a progressão das reservas apresentadas neste caso é mais consistente, embora no primeiro ano de ocorrência (2003) as reservas sejam inferiores às provisionadas pela empresa. Nos anos seguintes a tendência é de aumentar as reservas de tal forma que faz com que o valor total esteja mais equilibrado com o valor total da seguradora.

Desta forma, ao escolher este caso a empresa teria de aumentar as suas provisões para sinistros em 4 225 milhares de euros.

8. CONCLUSÕES

As Companhias de Seguros devem garantir uma gestão própria e qualificada de provisões técnicas para assegurar as responsabilidades que têm para com os seus clientes. Deste modo, os atuários assumem como uma das suas funções o provisionamento de estimativas para sinistros, sendo uma componente fulcral para a solvabilidade e para as contas de uma Seguradora.

O principal objetivo desta Dissertação foi o cálculo e análise das provisões para sinistros dos Danos Corporais na linha de negócio Automóvel de uma seguradora e o impacto que o fator cauda pode ter nas suas reservas. Para tal foram adotados quatro métodos, utilizando por base o modelo estocástico obtido por *Mack T.* (1993), na medida em que para responder às necessidades de obter os fluxos de caixa anuais, foram considerados o valor provisionado caso a caso pela Companhia de Seguros (Caso 2), um modelo de ajustamento Log-Linear de extrapolação (Caso 3) e um modelo referido por *Neuhaus W.* (2008) (Caso 4).

De acordo com os resultados obtidos através dos métodos aplicados neste estudo, a Companhia de Seguros apresenta suficiência de provisões em dois dos casos em estudo (Caso 1 e Caso 3) e insuficiência de provisões nos outros dois casos em estudo (Caso 2 e Caso 4), concluindo que não está a provisionar adequadamente.

No momento de optar por um destes métodos, é importante referir que à exceção do primeiro método, por ser um caso utilizado apenas academicamente dado que não considera o fator cauda, numa linha de negócio que apresente um desenvolvimento longo, todos os outros métodos são métodos válidos a aplicar.

Desta forma, com base nos resultados, na sua discussão e também na experiência do atuário, a escolha do melhor método para a Seguradora utilizar seria o Caso 4. A constante δ pode ser determinada atendendo à linha de negócio e considerando a experiência do atuário.

Além destes, existem outros tipos de métodos/modelos que também podiam ter sido adotados neste estudo, que refletem o impacto através da utilização da cauda de um triângulo de sinistros para a análise das reservas de uma seguradora, como os métodos de *Martin Bondy* e os Métodos Algébricos já referidos na secção 2 (*Halpert et al.*, 2013).

Ao fazer o provisionamento das reservas, a Companhia de Seguros calcula as provisões caso a caso, conhecendo bem os seus sinistros e tendo acesso a um historial de custo médio. Deste modo, ao analisar estes dados de forma agregada quer-se verificar se o valor corresponde ao que a empresa tem provisionado. Também existem vários fatores que podem influenciar tais como determinados anos em que ocorreram sinistros com situações atípicas.

Posto isto, uma das principais limitações que o presente estudo tem é a falta de informação para poder avaliar se o valor que é provisionado é adequado ou não. Contudo, e tal como mencionado, as empresas de seguros estão a par de toda a informação que engloba a ocorrência de um sinistro e, portanto, devem ter isso em conta de forma a tomarem melhor as suas decisões.

9. BIBLIOGRAFIA

- Autoridade de Supervisão de Seguros e Fundos de Pensões (ASF). (2004). *Circular nº28/2004*. Obtido de Autoridade de Supervisão de Seguros e Fundos de Pensões (ASF).
- Autoridade de Supervisão de Seguros e Fundos de Pensões (ASF). (2018). *Dados Estatísticos de Seguros 2018*. Autoridade de Supervisão de Seguros e Fundos de Pensões (ASF) .
- Borginho, H. M. (2001). Metodologias Estocásticas de Estimação de Provisões para Sinistros: Relatório de Estágio Curricular. *Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa*.
- Borginho, H. M. (2003). Provisões para Sinistros Não Vida - Metodologias de Estimação. *Departamento de Análises de Riscos e Solvência do ISP*.
- Carrato A., e. I. (2019). Claims Reserving with R: ChainLadder-0.2.10 Package Vignette. *CRAN Mirrors*.
- Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife, (pp. 1-26). *The Annals of Statistics*.
- Halpert A., e. I. (2013). The Estimation of Loss Development Tail Factors: A Summary Report. *Casualty Actuarial Society E-forum*.
- Kärikä M. (n.d.). Non-Life Insurance Mathematics (MTMS.02.053) Lecture Notes, (pp: 65-80). *University of Tartu, Faculty of Science and Technology, Institute of Mathematics and Statistics* .
- Mack, T. (1993). *Distribution-Free Calculation of the Standard Error of Chain Ladder Reserve Estimates*, Vol. 23 nº2 (pp: 213-225). Munich Re, Munich: ASTIN Bulletin.
- Mack, T. (1999). *The Standard Error of Chain Ladder Reserve Estimates: Recursive Calculation and Inclusion of a Tail Factor*, Vol. 23, nº2, (pp: 361-366). Munich Re, Munich: ASTIN Bulletin.
- Neuhaus, W. (2008). Lecture Notes on Estimating Outstanding Claims in General Insurance.
- Norberg, R. (1986). A Contribution to Modelling of IBRN Claims, (pp: 155-203). *Scandinavian Actuarial Journal*.
- Rohatgi, V. K. (2015). *An Introduction to Probability and Statistics*. Obtido de John Wiley & Sons.
- Straubs E., G. D. (1998). The Faculty and Institute of Actuaries Claims Reserving Manual. Volume 1 and 2. *ASTIM Bulletin: The Journal of IAA*.
- Taylor, G. (2000). *Loss Reserving: An Actuarial Perspective*. Boston: Kluwer Academic Publisher.

